



SEQUENCE LISTING

<10> Von der Muelbe, Florian
 Hoerr, Ingmar
 Pascolo, Steve

<120> Pharmaceutical composition containing a stabilised mRNA
 optimised for translation in its coding regions

<130> 2793-1-001PCT/CIP

<140> 10/729,830

<141> 2003-12-05

<150> PCT/EP02/06180

<151> 2002-06-05

<160> 13

<170> PatentIn Ver. 2.1

<210> 1

<211> 774

<212> DNA

<213> Influenza virus

<220>

<223> Influenza matrix: wildtype gene (for comparison)

<220>

<223> Start codon: atg (nucleotides 11 to 13), stop
 codon: tga (nucleotides 767 to 769)

<400> 1

agatctaaag atgagtcttc taaccgaggt cgaaacgtac gttctctcta tcatcccgtc 6
 0

aggccccctc aaagccgaga tcgcacagag acttgaagat gtctttgcag ggaagaacac 1
 20

cgatcttgag gttctcatgg aatggctaaa gacaagacca atcctgtcac ctctgactaa 1
 80

ggggatttta ggatttgtgt tcacgctcac cgtgcccagt gagcgaggac tgcagcgtag 2
 40

acgctttgtc caaaatgccc ttaatgggaa cggggatcca aataacatgg acaaagcagt 3
 00

taaactgtat aggaagctca agagggagat aacattccat ggggccaaag aaatctcact 3
 60

cagttattct gctgggtgcac ttgccagttg tatgggcctc atatacaaca ggatgggggc 4
 20

tgtgaccact gaagtggcat ttggcctggg atgtgcaacc tgtgaacaga ttgctgactc 4
 80

ccagcatcgg tctcataggg aaatgggtgac aacaaccaac ccactaatca gacatgagaa 5

40

cagaatgggtt ttagccagca ctacagctaa ggctatggag caaatggctg gatcgagtga 6

00

gcaagcagca gaggccatgg aggttgctag tcaggctagg caaatgggtgc aagcgatgag 6

60

aaccattggg actcatccta gctccagtgc tggctctgaaa aatgatcttc ttgaaaattt 7

20

gcaggcctat cagaaacgaa tgggggtgca gatgcaacgg ttcaagtga ctag 7

74

<210> 2

<211> 252

<212> PRT

<213> Influenza virus

<400> 2

Met Ser Leu Leu Thr Glu Val Glu Thr Tyr Val Leu Ser Ile Ile Pro
1 5 10 15Ser Gly Pro Leu Lys Ala Glu Ile Ala Gln Arg Leu Glu Asp Val Phe
20 25 30Ala Gly Lys Asn Thr Asp Leu Glu Val Leu Met Glu Trp Leu Lys Thr
35 40 45Arg Pro Ile Leu Ser Pro Leu Thr Lys Gly Ile Leu Gly Phe Val Phe
50 55 60Thr Leu Thr Val Pro Ser Glu Arg Gly Leu Gln Arg Arg Arg Phe Val
65 70 75 80Gln Asn Ala Leu Asn Gly Asn Gly Asp Pro Asn Asn Met Asp Lys Ala
85 90 95Val Lys Leu Tyr Arg Lys Leu Lys Arg Glu Ile Thr Phe His Gly Ala
100 105 110Lys Glu Ile Ser Leu Ser Tyr Ser Ala Gly Ala Leu Ala Ser Cys Met
115 120 125Gly Leu Ile Tyr Asn Arg Met Gly Ala Val Thr Thr Glu Val Ala Phe
130 135 140Gly Leu Val Cys Ala Thr Cys Glu Gln Ile Ala Asp Ser Gln His Arg
145 150 155 160Ser His Arg Gln Met Val Thr Thr Thr Asn Pro Leu Ile Arg His Glu
165 170 175Asn Arg Met Val Leu Ala Ser Thr Thr Ala Lys Ala Met Glu Gln Met
180 185 190

Ala Gly Ser Ser Glu Gln Ala Ala Glu Ala Met Glu Val Ala Ser Gln
 195 200 205

Ala Arg Gln Met Val Gln Ala Met Arg Thr Ile Gly Thr His Pro Ser
 210 215 220

Ser Ser Ala Gly Leu Lys Asn Asp Leu Leu Glu Asn Leu Gln Ala Tyr
 225 230 235 240

Gln Lys Arg Met Gly Val Gln Met Gln Arg Phe Lys
 245 250

<210> 3

<211> 775

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: Influenza
 matrix: gene with increased G/C-content

<220>

<223> Start codon: atg (nucleotides 11 to 13), stop
 codon: tga (nucleotides 767 to 769)

<400> 3

agatctaaag atgagcctgc tgaccgaggt ggagacctac gtgctgagca tcatccccag 6
 0
 cggccccctg aaggccgaga tcgcccagag gctggaggac gtgttcgccg gcaagaacac 1
 20
 cgacctggag gtgctgatgg agtggctgaa gaccaggccc atcctgagcc ccctgaccaa 1
 80
 gggcatcctg ggcttcgtgt tcaccctgac cgtgcccagc gagcgcgggc tgcagcgccg 2
 40
 ccgcttcgtg cagaacgccc tgaacggcaa cggcgacccc aacaacatgg acaaggccgt 3
 00
 gaagctgtac aggaagctga agagggagat caccttccac ggcgccaagg agatcagcct 3
 60
 gagctacagc gccggcgccc tggccagctg catgggcctg atctacaaca ggatgggcgc 4
 20
 cgtgaccacc gaggtggcct tcggcctggt gtgcgccacc tgcgagcaga tcgccgacag 4
 80
 ccagcaccgc agccacaggc agatggtgac caccaccaac ccctgatca ggcacgagaa 5
 40
 caggatggtg ctggccagca ccaccgcaa ggccatggag cagatggccg gcagcagcga 6
 00
 gcaggccgcc gaggccatgg aggtggccag ccaggccagg cagatggtgc aggccatgag 6
 60
 gaccatcggc acccacccca gcagcagcgc cggcctgaag aacgacctgc tggagaacct 7
 20

gcaggcctac cagaagcgca tgggcgtgca gatgcagcgc ttcaagtga ctagt 7
75

<210> 4

<211> 844

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: Influenza
matrix: gene for secreted form (with N-terminal
signal sequence) with increased G/C-content

<220>

<223> Start codon: atg (nucleotides 11 to 13), stop
codon: tga (nucleotides 836 to 838)

<400> 4

agatctaaag atggccgtca tggccccccg caccctggtg ctgctgctga gcggcgccct 6
0
ggccctgacc cagacctggg ctagcctgct gaccgaggtg gagacctacg tgctgagcat 1
20
catccccagc ggccccctga aggccgagat cgcccagagg ctggaggacg tgttcgccgg 1
80
caagaacacc gacctggagg tgctgatgga gtggctgaag accaggccca tctgagccc 2
40
cctgaccaag ggcattcctgg gcttcgtgtt caccctgacc gtgccagcgc agcgcggcct 3
00
gcagcgccgc cgcttcgtgc agaacgccct gaacggcaac ggcgacccca acaacatgga 3
60
caaggccgtg aagctgtaca ggaagctgaa gagggagatc accttccacg gcgccaagga 4
20
gatcagcctg agctacagcg ccggcgccct ggccagctgc atgggcctga tctacaacag 4
80
gatgggcgcc gtgaccaccg aggtggcctt cggcctggtg tgcgccacct gcgagcagat 5
40
cgccgacagc cagcaccgca gccacaggca gatggtgacc accaccaacc ccctgatcag 6
00
gcacgagAAC aggatggtgc tggccagcac caccgccaag gccatggagc agatggccgg 6
60
cagcagcgag caggccgccc aggccatgga ggtggccagc caggccaggc agatggtgca 7
20
ggccatgagg accatcggca cccaccccag cagcagcgcc ggccctgaaga acgacctgct 7
80
ggagaacctg caggcctacc agaagcgcat gggcgtgcag atgcagcgct tcaagtgaac 8
40
tagt 8
44

<210> 5

<211> 942

<212> RNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: Influenza
matrix: mRNA with stabilisation sequences

<220>

<223> The stabilisation sequences are derived from 5'-
and 3'-UTRs of β -globin-mRNA from *Xenopus laevis*,
respectively.

<220>

<223> Start codon: aug (nucleotides 56 to 58), stop
codon: uga (nucleotides 812 to 814)

<400> 5

```

gcuuguucuu uuugcagaag cucagaauaa acgcucaacu uuggcagauc uaaagaugag 6
0
ucuucuaacc gaggucgaaa cguacguucu cucuaucauc ccgucaggcc cccucaaagc 1
20
cgagaucgca cagagacuug aagaugucuu ugcaggggaag aacaccgauc uugagguucu 1
80
cauggaaugg cuaaagacaa gaccaauccu gucaccucug acuaagggga uuuuaggauu 2
40
uguguucacg cucaccgugc ccagugagcg aggacugcag cguagacgcu uuguccaaaa 3
00
ugcccuuaau gggaacgggg auctaaauaa cauggacaaa gcaguuaaac uguauaggaa 3
60
gcucaagagg gagauaacau uccauggggc caaagaauc ucacucaguu auucugcugg 4
20
ugcacuugcc aguuguaugg gccucauaua caacagggaug ggggcuguga ccacugaagu 4
80
ggcauuuggc cugguaugug caaccuguga acagauugcu gacucccagc aucggucuca 5
40
uaggcaaaug gugacaacaa ccaaccacu aaucagacau gagaacagaa ugguuuuagc 6
00
cagcacuaca gcuaaggcua uggagcaaa uggcuggaucg agugagcaag cagcagaggc 6
60
cauggagguu gcuagucagg cuaggcaaa uggugcaagcg augagaacca uugggacuca 7
20
uccuagcucc agugcugguc ugaaaaauga ucuucuugaa aaauugcagg ccuaucagaa 7
80
acgaaugggg gugcagaugc aacgguucaa gugaacuagu gacugacuag cccgcugggc 8
40
cucccaacgg gccuccucc ccuccuugca ccaaaaaaaaa aaaaaaaaaa aaaaaaaaaa 9
00
aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aa
42

```

<210> 6

<211> 942

<212> RNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: Influenza
matrix: mRNA with increased G/C-content and
stabilisation sequences

<220>

<223> The stabilisation sequences are derived from 5'-
and 3'-UTRs of β -globin-mRNA from *Xenopus laevis*,
respectively.

<220>

<223> Start codon: aug (nucleotides 56 to 58), stop
codon: uga (nucleotides 812 to 814)

<400> 6

```

gcuuguucuu uuugcagaag cucagaauaa acgcucaacu uuggcagauc uaaagaugag 6
0
ccugcugacc gagguggaga ccuacgugcu gagcaucauc cccagcggcc cccugaaggc 1
20
cgagaucgcc cagaggcugg aggacguguu cgccggcaag aacaccgacc uggaggugcu 1
80
gauggagugg cugaagacca ggcccauccu gagccccug accaagggca uccugggcuu 2
40
cguguucacc cugaccgugc ccagcgagcg cggccugcag cgccgccgcu ucgugcagaa 3
00
cgcccugaac ggcaacggcg accccaacaa cauggacaag gccgugaagc uguacaggaa 3
60
gcugaagagg gagaucaccu uccacggcgc caaggagauc agccugagcu acagcgccgg 4
20
cgcccuggcc agcugcaugg gccugaucua caacaggau ggcgccguga ccaccgaggu 4
80
ggccuucggc cuggugugcg ccaccugcga gcagaucgcc gacagccagc accgcagcca 5
40
caggcagaug gugaccacca ccaacccccu gaucaggcac gagaacagga uggugcuggc 6
00
cagcaccacc gccaaggcca uggagcagau ggccggcagc agcgagcagg ccgccgaggc 6
60
cauggaggug gccagccagg ccaggcagau ggugcaggcc augaggacca ucggcaccca 7
20
ccccagcagc agcgccggcc ugaagaacga ccugcuggag aaccugcagg ccuaccagaa 7
80
gcgcaugggc gugcagaugc agcgcuucaa gugaacuagu gacugacuag cccgcugggc 8
40
cucccaacgg gccuccucc ccuccuugca ccaaaaaaaaa aaaaaaaaaa aaaaaaaaaa 9
00
aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aa
42

```

<210> 7

<211> 1011

<212> RNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: Influenza
matrix: mRNA coding for secreted form with
increased G/C-content and stabilisation sequences

<220>

<223> Start codon: aug (nucleotides 56 to 58), stop
codon: uga (nucleotides 881 to 883)

<400> 7

```

gcuuguucuu uuugcagaag cucagaauaa acgcucaacu uuggcagauc uaaagauggc 6
0
cgucauggcc ccccgacacc uggugcugcu gcugagcggc gcccuggccc ugacccagac 1
20
cugggccagc cugcugaccg agguggagac cuacgugcug agcaucaucc ccagcggccc 1
80
ccugaaggcc gagaucgccc agaggcugga ggacguguuc gccggcaaga acaccgaccu 2
40
ggaggugcug auggaguggc ugaagaccag gcccauccug agccccuga ccaagggcau 3
00
ccugggcuuc guguucaccc ugaccgugcc cagcgagcgc ggccugcagc gccgccgcuu 3
60
cgugcagaac gcccugaacg gcaacggcga cccaacaac auggacaagg ccgugaagcu 4
20
guacaggaag cugaagaggg agaucaccuu ccacggcgcc aaggagauca gccugagcua 4
80
cagcgccggc gcccuggcca gcugcauggg ccugaucuaa aacaggauagg gcgccgugac 5
40
caccgaggug gccuucggcc uggugugcgc caccugcgag cagaucgccg acagccagca 6
00
ccgcagccac aggcagaugg ugaccaccac caacccccug aucaggcacg agaacaggau 6
60
ggugcuggcc agcaccaccg ccaaggccau ggagcagaug gccggcagca gcgagcaggc 7
20
cgccgaggcc auggaggugg ccagccaggc caggcagaug gugcaggcca ugaggaccu 7
80
cggcacccac cccagcagca gcgccggccu gaagaacgac cugcuggaga accugcaggc 8
40
cuaccagaag cgcauggggc ugcagaugca gcgcuucaag ugaacuagug acugacuagc 9
00
ccgcugggcc uccaacggg ccuccuccc cuccuugcac caaaaaaaaa aaaaaaaaaa 9
60
aaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa a          1
011

```

<210> 8
 <211> 940
 <212> DNA
 <213> Homo sapiens

<220>
 <223> MAGE1: wildtype-gene (for comparison)

<220>
 <223> Start codon: atg (nucleotides 5 to 7), stop codon:
 tga (nucleotides 932 to 934)

<400> 8
 catcatgtct cttgagcaga ggagtctgca ctgcaagcct gaggaagccc ttgaggccca 6
 0
 acaagaggcc ctgggcctgg tgtgtgtgca ggctgccacc tcctcctcct ctctcttggt 1
 20
 cctgggcacc ctggaggagg tgcccactgc tgggtcaaca gatcctcccc agagtcctca 1
 80
 gggagcctcc gcctttccca ctaccatcaa cttcactcga cagaggcaac ccagtgaggg 2
 40
 ttccagcagc cgtgaagagg aggggccaaag cacctcttgt atcctggagt ccttggtccg 3
 00
 agcagtaatc actaagaagg tggctgattt ggttggtttt ctgctcctca aatatcgagc 3
 60
 cagggagcca gtcacaaagg cagaaatgct ggagagtgtc atcaaaaatt acaagcactg 4
 20
 ttttcctgag atcttcggca aagcctctga gtccttgag ctggtccttg gcattgacgt 4
 80
 gaaggaagca gacccaccg gccactccta tgtccttgtc acctgcctag gtctctccta 5
 40
 tgatggcctg ctgggtgata atcagatcat gcccaagaca ggcttcctga taattgtcct 6
 00
 ggtcatgatt gcaatggagg gcggccatgc tcctgaggag gaaatctggg aggagctgag 6
 60
 tgtgatggag gtgtatgatg ggagggagca cagtgcctat ggggagccca ggaagctgct 7
 20
 cacccaagat ttggtgcagg aaaagtacct ggagtaccgg caggtgccgg acagtgatcc 7
 80
 cgcacgctat gagttcctgt ggggtccaag ggccctcgct gaaaccagct atgtgaaagt 8
 40
 ccttgagtat gtgatcaagg tcagtgcaag agttcgcttt ttcttcccat ccctgcgatga 9
 00
 agcagctttg agagaggagg aagagggagt ctgagcatga 9
 40

<210> 9
 <211> 308
 <212> PRT
 <213> Homo sapiens

<220>

<223> Tumor antigen MAGE1: protein sequence

<400> 9

Ser	Leu	Glu	Gln	Arg	Ser	Leu	His	Cys	Lys	Pro	Glu	Glu	Ala	Leu	Glu	1	5	10	15
Ala	Gln	Gln	Glu	Ala	Leu	Gly	Leu	Val	Cys	Val	Gln	Ala	Ala	Thr	Ser	20	25	30	
Ser	Ser	Ser	Pro	Leu	Val	Leu	Gly	Thr	Leu	Glu	Glu	Val	Pro	Thr	Ala	35	40	45	
Gly	Ser	Thr	Asp	Pro	Pro	Gln	Ser	Pro	Gln	Gly	Ala	Ser	Ala	Phe	Pro	50	55	60	
Thr	Thr	Ile	Asn	Phe	Thr	Arg	Gln	Arg	Gln	Pro	Ser	Glu	Gly	Ser	Ser	65	70	75	80
Ser	Arg	Glu	Glu	Glu	Gly	Pro	Ser	Thr	Ser	Cys	Ile	Leu	Glu	Ser	Leu	85	90	95	
Phe	Arg	Ala	Val	Ile	Thr	Lys	Lys	Val	Ala	Asp	Leu	Val	Gly	Phe	Leu	100	105	110	
Leu	Leu	Lys	Tyr	Arg	Ala	Arg	Glu	Pro	Val	Thr	Lys	Ala	Glu	Met	Leu	115	120	125	
Glu	Ser	Val	Ile	Lys	Asn	Tyr	Lys	His	Cys	Phe	Pro	Glu	Ile	Phe	Gly	130	135	140	
Lys	Ala	Ser	Glu	Ser	Leu	Gln	Leu	Val	Phe	Gly	Ile	Asp	Val	Lys	Glu	145	150	155	160
Ala	Asp	Pro	Thr	Gly	His	Ser	Tyr	Val	Leu	Val	Thr	Cys	Leu	Gly	Leu	165	170	175	
Ser	Tyr	Asp	Gly	Leu	Leu	Gly	Asp	Asn	Gln	Ile	Met	Pro	Lys	Thr	Gly	180	185	190	
Phe	Leu	Ile	Ile	Val	Leu	Val	Met	Ile	Ala	Met	Glu	Gly	Gly	His	Ala	195	200	205	
Pro	Glu	Glu	Glu	Ile	Trp	Glu	Glu	Leu	Ser	Val	Met	Glu	Val	Tyr	Asp	210	215	220	
Gly	Arg	Glu	His	Ser	Ala	Tyr	Gly	Glu	Pro	Arg	Lys	Leu	Leu	Thr	Gln	225	230	235	240
Asp	Leu	Val	Gln	Glu	Lys	Tyr	Leu	Glu	Tyr	Arg	Gln	Val	Pro	Asp	Ser	245	250	255	

3

4

```

caggaccugg uccaggagaa guaccuggag uaccgccagg ucccggacag cgacccggcg 7
80
cgcuacgagu uccugugggg cccgcgcgcg cuggcggaga cgagcuacgu caagguccug 8
40
gaguacguca ucaaggucag cgcgcgcguc cgcuuuuuu ucccagagccu gcgcgaggcg 9
00
gcgcgcgcgc aggaggagga gggcgucuga gcgugauga 9
39

```

```

<210> 11
<211> 939
<212> RNA
<213> Artificial Sequence

```

```

<220>
<223> Description of Artificial Sequence: MAGE1: mRNA
      with alternative use of codon

```

```

<220>
<223> Start codon: aug (nucleotides 1 to 3), stop codon:
      uga (nucleotides 937 to 939)

```

```

<400> 11
augagccugg agcagcgcag ccugcacugc aagcccaggagg aggccugga ggcccagcag 6
0
gaggcccugg gccuggugug cgugcaggcc gccaccagca gcagcagccc ccuggugcug 1
20
ggcaccucugg aggaggugcc caccgccggc agcaccgacc ccccccagag ccccagggc 1
80
gccagcgcgu uccccaccac caucaacuuc acccgccagc gccagcccag cgagggcagc 2
40
agcagccgcg aggaggaggg cccagcacc agcugcaucc uggagagccu guuccgcgcc 3
00
gugaucacca agaagguggc cgaccuggug ggcuuccugc ugcugaagua ccgcgcccgc 3
60
gagcccguga ccaaggccga gaugcuggag agcgugauca agaacuacaa gcacugcuuc 4
20
cccgagaucu ucggcaaggc cagcgagagc cugcagcugg uguucggcau cgacgugaag 4
80
gaggccgacc ccaccggcca cagcuacgug cuggugaccu gccuggggccu gagcuacgac 5
40
ggccugcugg gcgacaacca gaucaugccc aagaccggcu uccugaucau cgugcuggug 6
00
augaucgcca uggagggcgcg ccacgcccc gagggaggaga ucugggagga gcugagcgug 6
60
auggaggugu acgacggccg cgagcacagc gccuacggcg agccccgcaa gcugcugacc 7
20
caggaccugg ugcaggagaa guaccuggag uaccgccagg ugcccagacag cgaccccgcc 7
80
cgcuacgagu uccugugggg ccccgcgccc cuggccgaga ccagcuacgu gaaggugcug 8
40

```

gaguacguga ucaaggugag cgcccgcgug cgcuucuucu uccccagccu gcgcgaggcc 9
 00
 gcccgcgcg aggaggagga gggcguguga gccugauga 9
 39

<210> 12
 <211> 7
 <212> RNA
 <213> Artificial Sequence

<220>
 <223> Description of Artificial Sequence: Sequence motif
 recognizable for an endonuclease contained in the
 3'UTR-segment of the gene coding for the
 transferrin receptor (see p.10 of description)

<400> 12
 gaacaag 7

<210> 13
 <211> 13
 <212> RNA
 <213> Artificial Sequence

<220>
 <223> Description of Artificial Sequence: Kozak
 sequence, ribosome binding site (see p. 12 of
 description)

<400> 13
 gccgccacca ugg 1
 3